A corrected proof of the stochastic verification theorem within the framework of viscosity solutions

Fausto Gozzi* Andrzej Święch† Xun Yu Zhou‡

March 23, 2006

Abstract

We present a full and corrected proof of the stochastic verification theorem that was first obtained by X. Y. Zhou, J. Yong and X. Li, *Stochastic Verification theorems within the framework of viscosity solutions*, SIAM J. Control Optim., 35 (1997), pp. 243–253.

Keywords. Stochastic optimal control, verification theorem, HJB equation, viscosity solution, superdifferential, Lebesgue point.

AMS(MOS) Subject Classifications. 93E20, 49L20, 49L25.

---

*Dipartimento di Matematica per le Decisioni, Economiche, Finanziarie e Assicurative, Università di Roma "La Sapienza", Via del Castro Laurenziano 9, 00161 Roma, Italy. <fausto.gozzi@uniroma1.it>. Supported by the "Progetto di Ateneo" (years 1999-2001) on "Methods of static and dynamic optimization in Economics and Finance", Università di Roma "La Sapienza".

†School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, U.S.A. <swiech@math.gatech.edu>. Supported by the NSF grant DMS 0098565.

‡Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong. <xyzhou@se.cuhk.edu.hk>. Supported by the RGC Earmarked Grants CUHK 4234/01E.

1 Introduction

Dynamic programming approach is applied to optimal control problems, deterministic and stochastic alike, primarily via the so-called verification theorem. The verification theorem verifies whether a given admissible control is optimal and, more importantly, suggests a way of constructing optimal feedback controls. The classical version of the verification theorem (see, e.g., [4]), as with the dynamic programming at large, has the inherent deficiency of having to assume the smoothness of the value function, whereas it is by now well known that even a simplest optimal control problem may not possess a smooth value function. Therefore, it is imperative, to both the theoretical development and practical applications, that a non-smooth version of the verification theorem be available. A verification theorem for deterministic optimal control, within the framework of one of the non-smooth analyses – namely viscosity solutions, was put forward by Zhou in [8]. The stochastic version of the result in [8] was first correctly stated in [9] however with an incorrect proof. Specifically, in a key step in the proof of [9] a test function for superdifferential was employed and Ito's formula was applied. However, the superdifferential is pointwisely associated with probability sample points, hence the test function also implicitly depends on the sample points. In this case Ito's formula would not apply. Another proof was presented in the book [7, Chapter 5, Section 5.2]. It fixed the aforementioned problem of the original proof of [9] by incorporating the conditional probability, but unfortunately it still contained some, albeit technically subtle, gaps (see Remark 4.2 below for details). In this note we give a corrected proof of the result in [9]. The proof presented here is based on arguments of the one in [7] and may at first glance appear to be very similar. However it has a few delicate differences that take care of the points that were missing in the proof of [7]. Since the verification theorem is a result of fundamental importance we think it is worth to present its full and detailed proof even if it may repeat some of the material available in the literature.

2 The stochastic optimal control problem

In this section we introduce a class of stochastic optimal control problems that we consider in the paper. We will use the following basic notation throughout the paper. A' will denote the transpose of a matrix A, and $\mathbb{R}^{n\times k}$ and $S^{n\times n}$ will denote respectively the set of all $n\times k$ matrices and the set of all $n\times n$ symmetric matrices. Given a probability space $(\Omega,\mathcal{F},P)$ with a filtration $\{\mathcal{F}_t : a\le t\le b\}$ $(-\infty < a < b < +\infty)$, a separable Banach space Z with norm $|\cdot|_Z$, and $1\le p<+\infty$, we define the set $L^p_{\mathcal{F}_t}(a,b;Z) = \{\phi(\cdot) = \{\phi(t,\omega) : a\le t\le b\} \mid \phi(\cdot)$ is an $\mathcal{F}_t$-adapted, Z-valued measurable process on $[a,b]$, and $\mathbb{E}\int_a^b |\phi(t,\omega)|_Z^p dt < +\infty\}$.

Let $T\in(0,+\infty)$ be a finite time horizon and let $(\Omega,\mathcal{F},\mathbb{P})$ be a complete probability space. For any initial data $s\in[0,T]$ and $y\in\mathbb{R}^n$, the state equation of the problem is

$$\begin{cases} dx(t) = b(t,x(t),u(t))dt + \sigma(t,x(t),u(t))dW(t), \\ x(s) = y, \end{cases} \tag{1}$$

where W is an $m-$dimensional standard Brownian motion defined on $(\Omega,\mathcal{F},\mathbb{P})$, $u(\cdot)$ (the control strategy) is a stochastic process with values in a control space U, adapted to the filtration (augmented by $\mathbb{P}$-null sets) generated by W. The functional to minimize is

$$J(s,y;u(\cdot)) = \mathbb{E}[\int_s^T f(t,x(t),u(t))dt + h(x(T))],$$

where $u(\cdot)$ is a given control strategy and $x(\cdot)$ is the corresponding state trajectory solving (1). The infimum will be taken over a suitable set of admissible controls that will be defined below. Throughout the paper we will always assume that b,σ,f and h satisfy the following hypothesis. (We use the symbol $|\cdot|$ to denote norms in various spaces.)

Hypothesis 2.1 *The functions*

$$\begin{aligned} b &: [0,T]\times\mathbb{R}^n\times U\to\mathbb{R}^n, \qquad \sigma:[0,T]\times\mathbb{R}^n\times U\to\mathbb{R}^{n\times m},\\ f &: [0,T]\times\mathbb{R}^n\times U\to\mathbb{R}, \qquad h:\mathbb{R}^n\to\mathbb{R}\end{aligned}$$

are uniformly continuous. Moreover there exists a constant L such that for functions $\phi = b,\sigma,f,h$ we have

(i) $|\phi(t,x_1,u)-\phi(t,x_2,u)|\le L|x_1-x_2|$, $\forall t\in[0,T]$, $\forall x_1,x_2\in\mathbb{R}^n$, $\forall u\in U$;

(ii) $|\phi(t,0,u)|\le L$.

We will work, as in [7, Chapter 5, Section 4] with the so-called weak formulation of control problems. For a given $s\in[0,T]$, we define the set of (weakly) admissible controls $U^w_{ad}[s,T]$ to be the collection of all 5-tuples $(\Omega,\mathcal{F},\mathbb{P},W,u(\cdot))$ satisfying the following conditions:

(i) $(\Omega,\mathcal{F},\mathbb{P})$ is a complete probability space;

(ii) $W=\{W(t): s\le t\le T\}$ is an $m-$dimensional standard Brownian motion defined on $(\Omega,\mathcal{F},\mathbb{P})$ over $[s,T]$ (with $W(s)=0$ almost surely), and $\mathcal{F}^s_t$ is the filtration generated by W augmented with all $\mathbb{P}$-null sets in $\mathcal{F}$.

(iii) $u:[s,T]\times\Omega\to U$, is a U-valued $\mathcal{F}^s_t$-adapted process where U is a given complete separable metric space.

We will write $(\Omega,\mathcal{F},\mathbb{P},W,u(\cdot))\in U^w_{ad}[s,T]$, but occasionally we will only write $u(\cdot)\in U^w_{ad}[s,T]$ if no ambiguity arises.

We remark that in light of Hypothesis 2.1 for every $(s,y)\in[0,T]\times\mathbb{R}^n$ and $(\Omega,\mathcal{F},\mathbb{P},W,u(\cdot))\in U^w_{ad}[s,T]$, there exists a unique solution $x(\cdot)$ of equation (1) (see for instance [7, Theorem 6.16]). It is called the state corresponding to the control strategy $u(\cdot)\in U^w_{ad}[s,T]$, and $(x(\cdot),u(\cdot))$ is called an admissible pair. The objective of the optimal control problem is to minimize the cost functional $J(s,y;u(\cdot))$ (for a given $(s,y)\in[0,T)\times\mathbb{R}^n$) over all $u(\cdot)\in U^w_{ad}[s,T]$. We denote the above problem by $C_{s,y}$ to emphasize the dependence on the initial time s and the initial state y. The value function for the problem is defined as

$$V(s,y)=\inf_{u(\cdot)\in U^w_{ad}[s,T]} J(s,y;u(\cdot)).$$

An admissible pair $(x^*(\cdot),u^*(\cdot))$ is called *optimal* for $C_{s,y}$ if $u^*(\cdot)$ achieves the minimum of $J(s,y;u(\cdot))$ over $U^w_{ad}[s,T]$.

In the dynamic programming approach to the problem one proves that V is the unique classical solution of an associated Hamilton–Jacobi–Bellman (HJB) equation if $V\in C^{1,2}([0,T]\times\mathbb{R}^n)$. Defining the current value Hamiltonian H_{CV} as

$$H_{CV}(t,x,q,Q;u)=\frac{1}{2}\mathrm{tr}(\sigma(t,x,u)'Q\sigma(t,x,u))+\langle q,b(t,x,u)\rangle+f(t,x,u),$$

and the infimum value Hamiltonian H_{MIN} as

$$H_{MIN}(t,x,q,Q)=\inf_{u\in U} H_{CV}(t,x,q,Q;u),$$

the HJB equation associated to the family of problems $\{C_{s,y}\}_{(s,y)\in[0,T]\times\mathbb{R}^n}$ is

$$\begin{cases} -v_t(t,x)-H_{MIN}\left(t,x,Dv(t,x),D^2v(t,x)\right)=0, & (t,x)\in[0,T)\times\mathbb{R}^n,\\ v(T,x)=h(x), & x\in\mathbb{R}^n, \end{cases} \tag{2}$$

where Dv and D^2v denote respectively the gradient and Hessian of v with respect to x.

As a part of the dynamic programming approach, the so-called verification technique plays an important role in testing for optimality of a given admissible pair, and (more importantly) in constructing optimal feedback controls. The classical verification theorem is as follows (see Fleming and Rishel [4, Theorem VI.4.1]):

Theorem 2.2 *Let $W \in C^{1,2}([0,T] \times \mathbb{R}^n)$ be a solution of the HJB equation (2). Then:*

(a) $W(s,y) \leq J(s,y;u(\cdot))$, for any $(s,y) \in [0,T] \times \mathbb{R}^n$ and any $u(\cdot) \in U^w_{ad}[s,T]$.

(b) Suppose a given admissible pair $(x^*(\cdot), u^*(\cdot))$ for the problem $C_{s,y}$ satisfies

$$-W_t\left(t, x^*\left(t\right)\right) - H_{CV}\left(t, x^*(t), DW\left(t, x^*(t)\right), D^2W\left(t, x^*(t)\right); u^*\left(t\right)\right) = 0 \tag{3}$$

$\mathbb{P}-a.s.$, a.e. $t \in [s,T]$, then $(x^*(\cdot), u^*(\cdot))$ is an optimal pair for the problem $C_{s,y}$.

Remark 2.3 Equality (3) is equivalent to the more familiar one

$$H_{CV}\left(t, x^*(t), DW\left(t, x^*(t)\right), D^2W\left(t, x^*(t)\right); u^*(t)\right) = H_{MIN}\left(t, x^*(t), DW\left(t, x^*(t)\right), D^2W\left(t, x^*(t)\right)\right).$$

However, as is well known in general the value function is not smooth and the HJB equation may not have smooth solution at all. The verification theorem whose proof is presented in this paper is a nonsmooth version of Theorem 2.2. We refer the reader to [9, 7] for more discussions on this verification theorem and its consequences.

3 Preliminaries: stochastic processes, Lebesgue points, and viscosity solutions

In this subsection we present some preliminary results needed in the main proof.

3.1 Viscosity solutions and semidifferentials

To make the paper self-contained we present the definition of viscosity solution we will be using in this paper. It is slightly different from the typical and most commonly used one (see [1, 5, 7]).

First, by a parabolic neighborhood of a point (t_0, x_0) we mean the intersection of an open neighborhood of (t_0, x_0) with the set of all (t,x) such that $t \geq t_0$.

Definition 3.1 *An uppersemicontinuous (respectively, lowersemicontinuous) function v on $(0,T] \times \mathbb{R}^n$ is called a viscosity subsolution (respectively, supersolution) of (2) if*

$$v(T,x) \leq (\geq) h(x) \qquad \textit{for } x \in \mathbb{R}^n$$

and if whenever $v - \phi$ attains a local maximum (respectively, minimum) at (t_0, x_0) in a parabolic neighborhood of (t_0, x_0) for $\phi \in C^{1,2}(0,T) \times \mathbb{R}^n$ then

$$-\phi_t(t_0, x_0) - H_{MIN}(t_0, x_0, D\phi(t_0, x_0), D^2\phi(t_0, x_0)) \leq (\geq) 0. \tag{4}$$

A function v is called a viscosity solution of (2) if it is both a viscosity subsolution and a viscosity supersolution of (2).

The typical definition uses full (two-sided) neighborhoods instead of the parabolic (one-sided) ones. However it is well known that both definitions are equivalent. To see this suppose that v is a viscosity subsolution in the two-sided sense and that $v - \phi$ has a one-sided local maximum at (t_0, x_0) which can obviously be assumed to be strict. Then it is easy to see that for small $\mu > 0$ the functions $v - \phi - \frac{\mu}{t - t_0}$ have two-sided local maxima at points (t_μ, x_μ) such that $(t_\mu, x_\mu) \to (t_0, x_0)$ as $\mu \to 0$. Therefore from the two-sided definition we get

$$-\phi_t(t_\mu, x_\mu) - H_{MIN}(t_\mu, x_\mu, D\phi(t_\mu, x_\mu), D^2\phi(t_\mu, x_\mu)) \leq -\frac{\mu}{(t - t_\mu)^2} \leq 0$$

and letting $\mu \to 0$ we obtain (4).

Definition 3.1 is equivalent to the one in which test functions are replaced by parabolic semijets, i.e. second order parabolic sub- and superdifferentials (see [1]).

Definition 3.2 *We say that* $(p,q,Q) \in D^{1,2,+}_{t^+,x} v(t,x)$, *the second order one-sided parabolic superdifferential of* v *at* (t,x) *if for all* $y \in \mathbb{R}^n$, $s \geq t$

$$v(s,y) \leq v(t,x) + p(s-t) + \langle q, y-x \rangle + \frac{1}{2} \langle Q(y-x), y-x \rangle + o(s-t+|y-x|^2).$$

The second order one-sided parabolic subdifferential of v *at* (t,x), $D^{1,2,-}_{t^+,x} v$ *is defined by reversing the inequality, i.e.* $D^{1,2,-}_{t^+,x} v(t,x) = -D^{1,2,+}_{t^+,x}(-v)(t,x)$.

The equivalence of Definition 3.1 and the definition in which derivatives of test functions are replaced by elements of super- and subdifferentials is established with the help of a well known result we present below and whose proof can be found for instance in [7, Chapter 4, Lemma 5.4].

Lemma 3.3 *Let* v *be an uppersemicontinuous function on* $(0,T) \times \mathbb{R}^n$ *and let* $(t_0, x_0) \in (0,T) \times \mathbb{R}^n$. *Then* $(p,q,Q) \in D^{1,2,+}_{t^+,x} v(t_0,x_0)$ *if and only if there exists a function* $\phi \in C^{1,2}(0,T) \times \mathbb{R}^n$ *such that* $v - \phi$ *attains a strict global maximum at* (t_0,x_0) *relative to the set of* (t,x) *such that* $t \geq t_0$ *and*

$$(\phi(t_0,x_0), \phi_t(t_0,x_0), D\phi(t_0,x_0), D^2\phi(t_0,x_0)) = (v(t_0,x_0), p, q, Q). \tag{5}$$

Moreover, if v *has polynomial growth, i.e. if*

$$|v(t,x)| \leq C(1+|x|^k) \quad \text{for some } k \geq 1, \ (t,x) \in (0,T) \times \mathbb{R}^n \tag{6}$$

then ϕ *can be chosen so that* $\phi, \phi_t, D\phi, D^2\phi$ *satisfy (6) (with possibly different constants* C*).*

3.2 Lebesgue points

Here we recall and present a few results on Lebesgue points for functions with values in abstract spaces.

Definition 3.4 *Let* Z *be a Banach space and let* $z : [a,b] \to Z$ *be a measurable function that is Bochner integrable. We say that* t *is a right Lebesgue point of* z *if*

$$\lim_{h \to 0^+} \frac{1}{h} \int_t^{t+h} |z(r) - z(t)|_Z dr = 0.$$

We then have the following result, see for instance [2, Theorem 9, p.49].

Lemma 3.5 *Let* $z : [a,b] \to Z$ *be as in Definition 3.4. Then the set of right Lebesgue points of* z *is of full measure in* $[a,b]$.

The following simple lemma ensures that any process in $L^1_{\mathcal{F}_t}(a,b;\mathbb{R}^k)$ for some $k \geq 1$ can be regarded as Bochner integrable function $z : [a,b] \to L^1(\Omega;\mathbb{R}^k)$.

Lemma 3.6 *Any* $z \in L^1_{\mathcal{F}_t}(a,b;\mathbb{R}^k)$ *is Bochner integrable when regarded as a map from* $[a,b]$ *to* $L^1(\Omega;\mathbb{R}^k)$.

Proof. According to [3, p.92] if $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, Z is a separable Hilbert space, and $\mathcal{F}$ is countably generated apart from null sets, then the space $L^1(\Omega; Z)$ is separable. Since the filtration generated by a finite dimensional Wiener process is always countably generated (see for instance [6, Example 4.2.1, Section 1]), the space $L^1(\Omega;\mathbb{R}^k)$ is separable. Since

$$\int_a^b |z(t,\cdot)|_{L^1(\Omega;\mathbb{R}^k)} dt < +\infty$$

we only need to show that the function z is measurable as a map from $[a,b]$ to $L^1(\Omega;\mathbb{R}^k)$. By the separability of $L^1(\Omega;\mathbb{R}^k)$ the measurability is equivalent to the weak measurability so it is enough to show that for every function $w \in L^\infty(\Omega;\mathbb{R}^k)$ the map

$$t \to \mathbb{E}(z(t,\omega)\cdot w(\omega))$$

is Lebesgue measurable. But this is obvious from Fubini's theorem as

$$\mathbb{E}\int_a^b |z(t,\omega)\cdot w(\omega)|dt < +\infty.$$

■

The following technical result will be used in the proof of the verification theorem.

Proposition 3.7 *Let the standing assumptions of Section 2 be satisfied and let $(s,y) \in [0,T]\times\mathbb{R}^n$. Given any admissible pair $(x(\cdot),u(\cdot))$ for the control problem $C_{s,y}$, define the processes*

$$z_1(r) = b(r,x(r),u(r)), \quad z_2(r) = \sigma(r,x(r),u(r))\,\sigma(r,x(r),u(r))'.$$

Then

$$\lim_{h\to 0^+} \mathbb{E}\frac{1}{h}\int_t^{t+h} |z_i(r)-z_i(t)|dr \to 0 \quad as \quad h\to 0^+, \quad a.e.\ t\in[s,T], \quad i=1,2. \tag{7}$$

Proof. It is clear by Hypothesis 2.1 that $z_1, z_2 \in L^1_{\mathcal{F}_t}(s,T;Z)$ with $Z=\mathbb{R}^n$ and $Z=\mathbb{R}^{n\times n}$ respectively. Hence by Lemma 3.6 they are Bochner integrable when regarded as $L^1(\Omega;Z)$-valued maps. It then follows from Lemma 3.5 that the set of right Lebesgue points of z_1 and z_2 as maps from $[s,T]$ to $L^1(\Omega;Z)$ is of full measure in $[s,T]$ which gives the claim. ■

4 Verification theorem

We are going to prove the following theorem:

Theorem 4.1 *Let $U \in C((0,T]\times\mathbb{R}^n)$ be a viscosity subsolution of the HJB (2) satisfying (6) for some $k\geq 1$ and such that $U(T,x)=h(x)$. Then:*

(a) $U(s,y) \leq J(s,y;u(\cdot))$, for any $(s,y)\in(0,T]\times\mathbb{R}^n$ and any $u(\cdot)\in U^w_{ad}[s,T]$.

(b) Fix any $(s,y)\in(0,T)\times\mathbb{R}^n$. Let $(x^*(\cdot),u^*(\cdot))$ be an admissible pair for the problem $C_{s,y}$. Suppose that there exists $(p^*,q^*,Q^*)\in L^2_{\mathcal{F}_t}(s,T;\mathbb{R})\times L^2_{\mathcal{F}_t}(s,T;\mathbb{R}^n)\times L^2_{\mathcal{F}_t}(s,T;S^{m\times n})$ (where the filtration $\mathcal{F}_t$ is $(\mathcal{F}^s_t)_{t\in[s,T]}$) such that for a.e. $t\in[s,T]$,

$$(p^*(t),q^*(t),Q^*(t)) \in D^{1,2,+}_{t+,x}W(t,x^*(t)),\ \mathbb{P}-a.s. \tag{8}$$

and

$$\mathbb{E}\int_s^T \left[p^*(t)+H_{CV}(t,x^*(t),q^*(t),Q^*(t);u^*(t))\right]dt \leq 0. \tag{9}$$

Then $(x^*(\cdot),u^*(\cdot))$ is an optimal pair for the problem $C_{s,y}$.

Proof. Part (a) is obvious since $U\leq V$ in view of the standard comparison results for bounded viscosity sub- and supersolutions (see for instance [5, Theorem 9.1, Chapter V]). We just have to notice that if U satisfies (6) and K is big enough then the function

$$U_\delta(t,x) = U(t,x) - \delta e^{K(T-t)}|x|^{k+1}$$

is a viscosity subsolution of (2) for every $\delta > 0$ that is bounded from above. Similarly one can perturb V to a supersolution V_δ since V satisfies (6) with $k = 1$ under Hypothesis 2.1 (see [7, p.178, Proposition 3.1]). The comparison then follows from the comparison for U_δ and V_δ upon letting $\delta \to 0$.

It remains to prove part (b) of the theorem. To simplify the notation we set

$$f^*(t) = f(t, x^*(t), u^*(t)), h^*(t) = h(x^*(t)), b^*(t) = b(t, x^*(t), u^*(t)), \sigma^*(t) = \sigma(t, x^*(t), u^*(t)).$$

Fix $t_0 \in [s, T]$ such that (8) holds at t_0, and (7) holds at t_0 for $z_1(\cdot) = b^*(\cdot)$ and $z_2(\cdot) = \sigma^*(\cdot)\,\sigma^*(\cdot)'$. The set of such points is of full measure in $[s, T]$ by Proposition 3.7.

We now fix $\omega_0 \in \Omega$ such that the regular conditional probability given $\mathcal{F}^s_{t_0}$, denoted by $\mathbb{P}\left(\cdot|\mathcal{F}^s_{t_0}\right)(\omega_0)$ is well defined (see [7, Section 1.3, Chapter 1] for the definition and properties of conditional expectation and the regular conditional probability). We consider the probability space $\left(\Omega, \mathcal{F}, \mathbb{P}\left(\cdot|\mathcal{F}^s_{t_0}\right)(\omega_0)\right)$. In this probability space the random variables

$$x^*(t_0), p^*(t_0), q^*(t_0), Q^*(t_0)$$

are almost surely constant (see [7, Proposition 2.13, Chapter 1]) and are equal to

$$x^*(t_0, \omega_0), p^*(t_0, \omega_0), q^*(t_0, \omega_0), Q^*(t_0, \omega_0)$$

respectively. Denote $x_0 = x^*(t_0, \omega_0)$. We remark that in this probability space the Brownian motion W is still a standard Brownian motion however now $W(t_0) = W(t_0, \omega_0)$ almost surely. The space is now equipped with a new filtration $(\mathcal{F}^{t_0}_r)_{r \in [t_0, T]}$ and the control process $u^*(\cdot)$ is adapted to this new filtration (see [7, p. 179] for more on this). The main thing is that for $\mathbb{P}$ a.s. ω_0 the process $x^*(\cdot)$ is a solution of equation (1) on (t_0, T) in the new probability space $\left(\Omega, \mathcal{F}, \mathbb{P}\left(\cdot|\mathcal{F}^s_{t_0}\right)(\omega_0)\right)$ with the initial condition $x^*(t_0) = x^*(t_0, \omega_0)$.

We will write $\mathbb{E}_{\omega_0}$ to denote the expectation with respect to the measure $\mathbb{P}\left(\cdot|\mathcal{F}^s_{t_0}\right)(\omega_0)$ and will write $L^p_{\mathcal{F}_t, \omega_0}(t_0, T; Z)$ and $L^p_{\omega_0}(\Omega; Z)$ when these spaces are defined with respect to this measure.

Using Lemma 3.3 we now take a test function $\phi \in C^{1,2}((0, T) \times \mathbb{R}^n)$ such that

$$\phi(t, x) > U(t, x), \quad \text{for every } (t, x) \in [0, T] \times \mathbb{R}^n, \ (t, x) \neq (t_0, x_0), \tag{10}$$

$$\left(\phi(t_0, x_0), \phi_t(t_0, x_0), D\phi(t_0, x_0), D^2\phi(t_0, x_0)\right) = \left(U(t_0, x_0), p^*(t_0, \omega_0), q^*(t_0, \omega_0), Q^*(t_0, \omega_0)\right), \tag{11}$$

and such that $\phi, \phi_t, D\phi, D^2\phi$ are polynomially bounded, i.e. they satisfy (6). It is important to bear in mind that ϕ is a fixed deterministic function when (t_0, ω_0) is fixed.

We are going to apply Ito's formula to ϕ on the probability space $\left(\Omega, \mathcal{F}, \mathbb{P}\left(\cdot|\mathcal{F}^s_{t_0}\right)(\omega_0)\right)$. To do this we observe that the processes

$$\phi_t(r, x^*(r)), \quad \langle D\phi(r, x^*(r)), b^*(r)\rangle, \quad \frac{1}{2}\mathrm{tr}\left[\sigma^*(r)'\, D^2\phi(r, x^*(r))\,\sigma^*(r)\right]$$

belong to $L^1_{\mathcal{F}_t, \omega_0}(t_0, T; \mathbb{R})$ due to Hypothesis 2.1 on b and σ, to the estimate

$$\mathbb{E}_{\omega_0} \sup_{t_0 \le r \le T} |x^*(r)|^l \le K_T(1 + |x_0|^l), \text{ for every } l \ge 1, \tag{12}$$

(see for instance [7, Theorem 6.16, Chapter 1]), and to the polynomial growth of $\phi, \phi_t, D\phi, D^2\phi$. Moreover, by the same argument, the process

$$\sigma^*(r)'\,\phi_x(r, x^*(r))$$

belongs to $L^2_{\mathcal{F}_t, \omega_0}(t_0, T; \mathbb{R}^n)$. Therefore (see for instance [7, Section 5.3, Chapter 1]) we can apply Ito's formula to obtain that for any $h > 0$

$$\begin{aligned}
& \phi(t_0 + h, x^*(t_0 + h)) - \phi(t_0, x^*(t_0)) \\
= & \int_{t_0}^{t_0+h} \left[\phi_t(r, x^*(r)) + \langle D\phi(r, x^*(r)), b^*(r)\rangle + \frac{1}{2}\mathrm{tr}\left(\sigma^*(r)'\, D^2\phi(r, x^*(r))\,\sigma^*(r)\right)\right] dr \\
& + \int_{t_0}^{t_0+h} \langle D\phi(r, x^*(r)), \sigma^*(r)\, dW(r)\rangle.
\end{aligned}$$

In particular, taking the expected value $\mathbb{E}_{\omega_0}$, dividing both sides by h, and using (10), it follows that

$$\begin{aligned}
& \mathbb{E}_{\omega_0} \frac{1}{h} \left[U \left(t_0 + h, x^* \left(t_0 + h \right) \right) - U \left(t_0, x^* \left(t_0 \right) \right) \right] \\
\leq \; & \mathbb{E}_{\omega_0} \left[\frac{1}{h} \int_{t_0}^{t_0+h} \left[\phi_t \left(r, x^* \left(r \right) \right) + \left\langle D\phi \left(r, x^* \left(r \right) \right), b^* \left(r \right) \right\rangle + \frac{1}{2} \mathrm{tr} \left(\sigma^* \left(r \right)' D^2 \phi \left(r, x^* \left(r \right) \right) \sigma^* \left(r \right) \right) \right] dr \right]. \quad (13)
\end{aligned}$$

We want to pass to the limit as $h \to 0^+$ above. To this end we treat each term of (13) separately.

(I) First of all, thanks to the continuity of x^* and to the continuity of ϕ_t, we get

$$\frac{1}{h} \int_{t_0}^{t_0+h} \phi_t \left(r, x^* \left(r \right) \right) dr \to \phi_t \left(t_0, x^* \left(t_0 \right) \right) \qquad \mathbb{P} \left(\cdot | \mathcal{F}_{t_0}^s \right) \left(\omega_0 \right) - \text{a.s.}$$

Now, using estimate (12) and the polynomial growth of ϕ_t, we can apply the Dominated Convergence Theorem to obtain

$$\lim_{h \to 0^+} \mathbb{E}_{\omega_0} \left| \frac{1}{h} \int_{t_0}^{t_0+h} \phi_t \left(r, x^* \left(r \right) \right) dr - \phi_t \left(t_0, x^* \left(t_0 \right) \right) \right| = 0$$

(II) With regard to the second term of (13) we have

$$\begin{aligned}
& \mathbb{E}_{\omega_0} \frac{1}{h} \int_{t_0}^{t_0+h} \left\langle D\phi \left(r, x^* \left(r \right) \right), b^* \left(r \right) \right\rangle dr - \left\langle D\phi \left(t_0, x^* \left(t_0 \right) \right), b^* \left(t_0 \right) \right\rangle \\
= \; & \mathbb{E}_{\omega_0} \frac{1}{h} \int_{t_0}^{t_0+h} \left\langle D\phi \left(r, x^* \left(r \right) \right) - D\phi \left(t_0, x^* \left(t_0 \right) \right), b^* \left(r \right) \right\rangle dr \\
& + \mathbb{E}_{\omega_0} \frac{1}{h} \int_{t_0}^{t_0+h} \left\langle D\phi \left(t_0, x^* \left(t_0 \right) \right), b^* \left(r \right) - b^* \left(t_0 \right) \right\rangle dr. \qquad (14)
\end{aligned}$$

Now

$$\begin{aligned}
& \left| \mathbb{E}_{\omega_0} \frac{1}{h} \int_{t_0}^{t_0+h} \left\langle D\phi \left(r, x^* \left(r \right) \right) - D\phi \left(t_0, x^* \left(t_0 \right) \right), b^* \left(r \right) \right\rangle dr \right| \\
\leq \; & \mathbb{E}_{\omega_0} \frac{1}{h} \int_{t_0}^{t_0+h} \left| D\phi \left(r, x^* \left(r \right) \right) - D\phi \left(t_0, x^* \left(t_0 \right) \right) \right| \left| b^* \left(r \right) \right| dr \\
\leq \; & \mathbb{E}_{\omega_0} \left[\frac{1}{h} \int_{t_0}^{t_0+h} \left| D\phi \left(r, x^* \left(r \right) \right) - D\phi \left(t_0, x^* \left(t_0 \right) \right) \right|^2 dr \right]^{\frac{1}{2}} \left[\frac{1}{h} \int_{t_0}^{t_0+h} \left| b^* \left(r \right) \right|^2 dr \right]^{\frac{1}{2}} \\
\leq \; & \left[\mathbb{E}_{\omega_0} \frac{1}{h} \int_{t_0}^{t_0+h} \left| D\phi \left(r, x^* \left(r \right) \right) - D\phi \left(t_0, x^* \left(t_0 \right) \right) \right|^2 dr \right]^{\frac{1}{2}} \left[\mathbb{E}_{\omega_0} \frac{1}{h} \int_{t_0}^{t_0+h} \left| b^* \left(r \right) \right|^2 dr \right]^{\frac{1}{2}}
\end{aligned}$$

Using Hypothesis 2.1 and estimate (12) we have

$$\begin{aligned}
\mathbb{E}_{\omega_0} \frac{1}{h} \int_{t_0}^{t_0+h} \left| b^* \left(r \right) \right|^2 dr \; & \leq \; \frac{1}{h} \int_{t_0}^{t_0+h} 2L^2 \left(1 + \mathbb{E}_{\omega_0} \left| x^* \left(r \right) \right|^2 \right) dr \\
& \leq \; 2L^2 \left(1 + K_T \left(1 + \left| x_0 \right|^2 \right) \right), \qquad (15)
\end{aligned}$$

while arguing as in part **(II)** for ϕ_t we get

$$\lim_{h \to 0^+} \mathbb{E}_{\omega_0} \frac{1}{h} \int_{t_0}^{t_0+h} \left| D\phi \left(r, x^* \left(r \right) \right) - D\phi \left(t_0, x^* \left(t_0 \right) \right) \right| dr = 0. \qquad (16)$$

¿From (15) and (16) it follows that the first term of (14) goes to zero as $h \to 0^+$. For the second term of (14) we observe that

$$\left|\mathbb{E}_{\omega_0} \frac{1}{h} \int_{t_0}^{t_0+h} \left\langle D\phi\left(t_0, x^*\left(t_0\right)\right), b^*\left(r\right) - b^*\left(t_0\right)\right\rangle dr\right| \leq \left|D\phi\left(t_0, x_0\right)\right| \mathbb{E}_{\omega_0} \frac{1}{h} \int_{t_0}^{t_0+h} \left|b^*\left(r\right) - b^*\left(t_0\right)\right| dr$$

However by the choice of t_0

$$\begin{aligned} 0 = \lim_{h\to 0^+} \mathbb{E}\frac{1}{h} \int_{t_0}^{t_0+h} \left|b^*\left(r\right) - b^*\left(t_0\right)\right| dr &= \lim_{h\to 0^+} \mathbb{E}\left[\mathbb{E}\left[\frac{1}{h} \int_{t_0}^{t_0+h} \left|b^*\left(r\right) - b^*\left(t_0\right)\right| dr \middle| \mathcal{F}_{t_0}^s\right]\right] \\ &= \lim_{h\to 0^+} \mathbb{E}\left[\mathbb{E}_{\omega_0} \frac{1}{h} \int_{t_0}^{t_0+h} \left|b^*\left(r\right) - b^*\left(t_0\right)\right| dr\right]. \end{aligned}$$

This implies that $\mathbb{E}_{\omega_0} \frac{1}{h} \int_{t_0}^{t_0+h} \left|b^*\left(r\right) - b^*\left(t_0\right)\right| dr$ converges to 0 in $L^1(\Omega; \mathbb{R}^1)$ as $h \to 0^+$. Hence, there is a subsequence $h_l \to 0^+$ so that for $\mathbb{P}$- a.s. ω_0

$$\mathbb{E}_{\omega_0} \frac{1}{h} \int_{t_0}^{t_0+h} \left|b^*\left(r\right) - b^*\left(t_0\right)\right| dr \to 0.$$

This proves that the second term of (14) goes to zero as $h_l \to 0^+$.

(III) The third term of (13) is treated using the same ideas. Since

$$\operatorname{tr}\left(ABC\right) = \operatorname{tr}\left(BCA\right), \quad \text{and} \quad \operatorname{tr}\left(AB\right) - \operatorname{tr}\left(AC\right) = \operatorname{tr}A\left(B - C\right),$$

we have

$$\begin{aligned} &\mathbb{E}_{\omega_0}\left[\frac{1}{h}\int_{t_0}^{t_0+h}\left[\frac{1}{2}\operatorname{tr}\left(\sigma^*\left(r\right)' D^2\phi\left(r, x^*\left(r\right)\right)\sigma^*\left(r\right)\right) - \frac{1}{2}\operatorname{tr}\left(\sigma^*\left(t_0\right)' D^2\phi\left(t_0, x^*\left(t_0\right)\right)\sigma^*\left(t_0\right)\right)\right] dr\right] \\ = \; &\mathbb{E}_{\omega_0}\left[\frac{1}{h}\int_{t_0}^{t_0+h}\frac{1}{2}\operatorname{tr}\left(\left[D^2\phi\left(r, x^*\left(r\right)\right) - D^2\phi\left(t_0, x^*\left(t_0\right)\right)\right]\sigma^*\left(r\right)\sigma^*\left(r\right)'\right) dr\right] \\ &+\mathbb{E}_{\omega_0}\left[\frac{1}{h}\int_{t_0}^{t_0+h}\frac{1}{2}\operatorname{tr}\left(D^2\phi\left(t_0, x^*\left(t_0\right)\right)\left[\sigma^*\left(r\right)\sigma^*\left(r\right)' - \sigma^*\left(t_0\right)\sigma^*\left(t_0\right)'\right]\right) dr\right]. \end{aligned} \tag{17}$$

Now, employing the same arguments used to show that the right hand side of (14) goes to zero, we obtain that the right hand side of (17) goes to 0 as some subsequence $h_{l'} \to 0^+$.

Summing up we have proved that for any sequence $h \to 0^+$ one can find a subsequence $h_{l''} \to 0^+$ so that

$$\begin{aligned} &\lim_{h_{l''}\to 0^+} \mathbb{E}_{\omega_0}\left[\frac{1}{h_{l''}}\int_{t_0}^{t_0+h_{l''}}\left[\phi_t\left(r, x^*\left(r\right)\right) + \left\langle D\phi\left(r, x^*\left(r\right)\right), b^*\left(r\right)\right\rangle + \frac{1}{2}\operatorname{tr}\left(\sigma^*\left(r\right)' D^2\phi\left(r, x^*\left(r\right)\right)\sigma^*\left(r\right)\right)\right] dr\right] \\ &= \phi_t\left(t_0, x_0\right) + \left\langle D\phi\left(t_0, x_0\right), b^*\left(t_0\right)\right\rangle + \frac{1}{2}\operatorname{tr}\left(\sigma^*\left(t_0\right)' D^2\phi\left(t_0, x_0\right)\sigma^*\left(t_0\right)\right). \end{aligned}$$

Consequently, the above holds for any sequence $h \to 0^+$. Passing to the limsup in (13) and using (11), we get that

$$\begin{aligned} &\limsup_{h\to 0^+} \mathbb{E}_{\omega_0}\frac{1}{h}\left[U\left(t_0 + h, x^*\left(t_0 + h\right)\right) - U\left(t_0, x^*\left(t_0\right)\right)\right] \\ \leq \; &\phi_t\left(t_0, x_0\right) + \left\langle D\phi\left(t_0, x_0\right), b^*\left(t_0\right)\right\rangle + \frac{1}{2}\operatorname{tr}\left(\sigma^*\left(t_0\right)' D^2\phi\left(t_0, x_0\right)\sigma^*\left(t_0\right)\right) \\ = \; &p^*\left(t_0, \omega_0\right) + \left\langle q^*\left(t_0, \omega_0\right), b^*\left(t_0, \omega_0\right)\right\rangle + \frac{1}{2}\operatorname{tr}\left(\sigma^*\left(t_0, \omega_0\right)' Q^*\left(t_0, \omega_0\right)\sigma^*\left(t_0, \omega_0\right)\right). \end{aligned} \tag{18}$$

Now, using (18) and applying Fatou's Lemma we get

$$
\begin{aligned}
& \limsup_{h\to 0^+} \mathbb{E}\left[\frac{1}{h}\left[U\left(t_0+h, x^*\left(t_0+h\right)\right)-U\left(t_0, x^*\left(t_0\right)\right)\right]\right] \\
= & \limsup_{h\to 0^+} \mathbb{E}\left[\mathbb{E}_{\omega_0} \frac{1}{h}\left[U\left(t_0+h, x^*\left(t_0+h\right)\right)-U\left(t_0, x^*\left(t_0\right)\right)\right]\right] \\
\leq & \mathbb{E}\left[\limsup_{h\to 0^+}\left[\mathbb{E}_{\omega_0} \frac{1}{h}\left[U\left(t_0+h, x^*\left(t_0+h\right)\right)-U\left(t_0, x^*\left(t_0\right)\right)\right]\right]\right] \\
\leq & \mathbb{E}\left[p^*\left(t_0\right)+\left\langle q^*\left(t_0\right), b^*\left(t_0\right)\right\rangle+\frac{1}{2} \operatorname{tr}\left(\sigma^*\left(t_0\right)' Q^*\left(t_0\right) \sigma^*\left(t_0\right)\right)\right] \qquad (19)
\end{aligned}
$$

for a.e. $t_0 \in [s,T]$. The rest of the proof goes exactly as in [7, p.272]. We apply Lemma 5.2 of [7] to the function $g(t) = \mathbb{E}U(t, x^*(t))$, and use (19) and (9) to obtain

$$
\mathbb{E}U(T, x^*(T)) - U(s,y) \leq -\mathbb{E}\int_s^T f^*(t)dt
$$

which, since $U \leq V$ and $U(T, x^*(T)) = V(T, x^*(T)) = h(x^*(T))$, yields

$$
\mathbb{E}[\int_s^T f\left(t, x^*(t), u^*(t)\right) dt + h(x^*(T))] \leq V(s,y),
$$

which means that the control $u^*(\cdot)$ is optimal. ∎

Remark 4.2 Now we are in a position to discuss the (subtle) gap presented in the proof of the book [7]. On page 271 of [7], equation (5.12) was derived by considering a Lebesgue point t of the integrand in the displayed equation right before (5.12). However, the test function ϕ there depends, though implicitly, on the variable t. Hence it is not clear if such a t exists in the first place. In this paper, we have fixed this problem via the Lebesgue points of the problem data ($b^*(\cdot)$ and $\sigma^*(\cdot)\sigma^*(\cdot)'$) viewed as $L^1(\Omega; \mathbb{R}^n)$ or $L^1(\Omega; \mathbb{R}^{n\times n})$-valued functions, together with some delicate estimates.

Acknowledgment. Part of this research was carried out when the third author was visiting Scuola Normale Superiore, Italy. The hospitality of the host, Giuseppe Da Prato, is gratefully acknowledged. Moreover, the authors thank the editors and two reviewers for their helpful comments and suggestions on an earlier version of the paper.

References

[1] M. G. Crandall, H. Ishii and P. L. Lions, *A user's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc., 27 (1992), pp. 1–67.

[2] J. Diestel and J. J. Uhl, Jr. *Vector Measures*, Mathematical Surveys and Monographs, 15, American Mathematical Society, Providence, Rhode Island, 1977.

[3] H. Doob, *Measure Theory,* Springer-Verlag, New York, 1994.

[4] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control,* Springer-Verlag, New York, 1975.

[5] W. H. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions,* Applications of Mathematics, 25, Springer-Verlag, New York, 1993.

[6] D. Revuz, M, Yor, Continuous Martingales and Brownian Motion, Third edition, Springer-Verlag, New York, 1999.

[7] J. Yong, X. Y. Zhou, *Stochastic Control: Hamiltonian Systems and HJB Equations*, Springer-Verlag, New York, 1999.

[8] X. Y. Zhou, *Verification theorems within the framework of viscosity solutions*, J. Math. Anal. Appl., 177 (1993), pp. 208– 225.

[9] X. Y. Zhou, J. Yong and X. Li, *Stochastic Verification theorems within the framework of viscosity solutions*, SIAM J. Control Optim., 37 (1997), pp. 243–253.